

An interactive flower image recognition system

Tzu-Hsiang Hsu · Chang-Hsing Lee · Ling-Hwei Chen

Published online: 6 March 2010

© Springer Science+Business Media, LLC 2010

Abstract In this paper, we present an interactive system for recognizing flower images taken by digital cameras. The proposed system provides an interactive interface allowing each user to draw an appropriate bounding window that contains the interested flower region. Then, a flower boundary tracing method is developed to extract the flower region as accurately as possible. In addition to the color and shape features of the whole flower region, the color and shape features of the pistil/stamen area will also be used to represent the flower characteristics more precisely. Experiments conducted on two distinct databases consisting of 24 species and 102 species have shown that our proposed system outperforms other approaches in terms of the recognition rate.

Keywords Flower image recognition · Image segmentation

1 Introduction

There are about 250,000 named species of flowering plants in the world. Everyday, we can see many blooming flowers in the roadside, garden, park, mountain path, wild field, etc. Generally, experienced taxonomists or botanists can identify plants according to their flowers. However, most people do know nothing about these wild flowers, even their names. To know the names or characteristics of the plants, we usually have to consult flower guide books or browse any relevant web pages on the Internet through keywords searching. Typically, such a keyword searching approach is not practical for most people.

This research was supported in part by the National Science Council of R.O.C. under contract NSC-97-2221-E-009-137.

T.-H. Hsu · L.-H. Chen (✉)

Institute of Multimedia and Engineering, National Chiao Tung University, 1001 University Road, Hsinchu, Taiwan 300, Republic of China
e-mail: lhchen@cc.nctu.edu.tw

C.-H. Lee

Department of Computer Science and Information Engineering, Chung Hua University, Hsinchu, Taiwan 300, Republic of China

Since digital cameras have been widely used for most people, it would be very useful to identify the blooming plant based on the flower images taken by a digital camera. The first problem in a flower recognition system is how to accurately extract the flower region from a natural complex background. Once the flower region is segmented, effective color, shape, and texture features are extracted for further recognition purpose.

Saitoh and Kaneko [11] proposed an automatic method for recognizing wild flowers using a frontal flower image and a leaf image taken by a digital camera. To take the flower images and the leaf images, they first placed a black sheet under the flowers or leaves, which is inconvenient and laborious. To well separate the flower and the leaf from the background, they used k-means clustering algorithm to model the background region. A total of 17 features that describe the color and shape properties of the flower and the leaf images were extracted for flower recognition by using a neural network. A recognition rate of 95% was obtained for the recognition of 20 sets of flower and leaf images from 16 species. The main problem with this approach is that it is inconvenient to take the images.

Das et al. [3] proposed an approach to indexing flower patent images using the domain knowledge of flower colors and their spatial locations. Generally, the colors appear in the flower regions are rarely green, black, gray, or brown and the background colors are usually visible along the periphery of the image. An automatic iterative segmentation algorithm exploiting the domain knowledge was developed to isolate the flower region from the background. Only the colors in the flower region instead of all colors in the entire image were used to index similar flowers. The color features include color names and their relative proportions in the flower region. Their flower indexing system provided queries using natural language color names and an example image. A flower database consists of 300 images was tested to demonstrate the effectiveness of their proposed approach. However, using color information alone, without considering the shape features, cannot recognize flower images effectively.

Hong et al. [7] proposed a flower image retrieval method based on the features extracted from the region-of-interest (ROI), which corresponds to the flower region. A segmentation method was proposed to separate the flower regions from background using color clustering method and the domain knowledge similar to that proposed by Das et al. [3]. The color histogram, which represents the color distribution of the flower region, as well as two shape features were extracted to search similar flower images. These two shape features, the centroid-contour-distance (CCD) curve and the angle-code-histogram (ACH), were extracted to characterize the shape of the flower contour. CCD curve measures the distances from all contour points to the center of the flower region. For each contour point, the angle between two approximate lines starting from and ending at the point will be accumulated to form ACH. Experimental results on 885 flower images from 14 plant species have shown that their approach outperforms the method based on the global color histogram proposed by Swain and Ballard [14] and the method proposed by Das et al. [3]. The main problem with this approach is that the CCD curve and ACH will be greatly affected if some petals fall off, bend, curl, twist, etc.

Zou and Nagy [16] developed a model-based interactive flower recognition system based on the concept of Computer Assisted Visual InterActive Recognition (CAVIAR). In the training process, each training image was interactively segmented in order to extract the flower regions. Domain-specific rose-curve model was then employed to fit the silhouette of each flower region. Eight model parameters, including the petal number, the ratio of the outer radius to the inner radius, and the first three moments of the hue and saturation histograms of the pixels within the rose curve, were extracted to recognize flower images. In the recognition process, an initial rose curve of the test flower image was estimated and superimposed on the test flower image. The first three candidates were displayed according to the model parameters extracted from the initial rose curve. The user can accept one of the

recognition results or tries to interactively adjust the parameters of the rose-curve model with mouse operations. According to the adjustment, the system will re-compute the model parameters and re-rank the recognition results. Such an interactive process will repeat until the user accepts the recognition result. One major problem of this system is that too many user interactions have to be conducted to get high recognition accuracy.

Nilsback and Zisserman [10] developed a visual vocabulary that explicitly describes the various characteristics (color, shape, and texture) of flowers. First, each image is automatically segmented into foreground region (flower part) and background region using the contrast dependent prior Markov random field (MRF) cost function [1] and optimized using graph cuts. The *HSV* color values of all pixels in the training images were then divided into V_c clusters using k-means clustering algorithm. The number of clusters V_c is optimized on the dataset. Then, a color vocabulary is constructed by the set of cluster centers (visual words). As a result, each image is represented by a V_c -dimensional normalized frequency histogram of the set of visual words. To describe the shape of each petal, a rotation invariant descriptor called scale-invariant feature transform (SIFT) descriptor [8] was computed on a regular grid and optimized over three parameters: the grid spacing M , the radius R of the support region for SIFT computation, and the number of clusters. Vector quantization was then applied to get the visual words representing the petal shapes. The frequency histogram corresponding to the shape visual words was calculated to describe the shape characteristic. To model the characteristic patterns on different petals, texture features were computed by convolving the image with maximum response 8 (MR8) filter bank [15]. The performance was optimized over the size of the square support regions of the MR8 filters. A vocabulary was created by clustering the texture descriptors of all training images and the frequency histogram was obtained for each image. For each characteristic (color, shape, or texture), the distance between two images is evaluated by the χ^2 measure of their frequency histograms. To get better performance, they combined these three vocabularies into a joint flower vocabulary and obtained a joint frequency histogram. A weight vector associated with the joint frequency histogram was introduced to optimize the performance. Experimental results on a dataset of 1360 images from 17 flower species have shown that the combined vocabulary outperforms each of the individual ones. Typically, there are too many parameters need to be optimized to get high recognition rate.

Saitoh et al. [13] extended the route tracing method [12] to automatically extract the flower boundary under the assumption that the flower region is focused and the background is out of focus. The extended route tracing method is based on the Intelligent Scissor (IS) approach [9] which searches a route that minimizes the sum of local costs according to a number of manually selected points on the visually identified flower boundary. Instead of minimizing the sum of local costs, the extended route tracing method tried to minimize the average cost defined as the sum of local costs divided by the route length. Four shape features (the ratio of the route length to the sum of distances between the gravity center and all boundary points, the number of petals, the central moment, and the roundness) as well as six color features (the x and y coordinates and the proportions of flower pixels accumulated in the two largest color cells in the *HS* color space) were extracted to recognize flower images. Experiments were conducted on 600 images from 30 species with 20 images per species. The recognition rates were 90.7%, 97.7%, and 99.0% if the correct one is included in the top one, top two, and top three candidates, respectively. It is worth to note that the number of petals will change if some petals fell off or were occluded by others.

Cho and Chi [2] proposed a structure-based flower image recognition method. The genetic evolution algorithm with adaptive crossover and mutation operations was employed to tune the learning parameters of the Backpropagation Through Structures algorithm [5]. A

region-based binary tree representation whose nodes correspond to the regions of the flower image and links represent the relationships among regions was constructed to represent the flower image content. Experimental results showed that the structural representation of flower images can produce a promising performance for flower image recognition in terms of generalization and noise robustness. In fact, the classification accuracy of the system depends on the selection of the feature values.

Fukuda et al. [4] developed a flower image retrieval system by combining multiple classifiers using fuzzy c-means clustering algorithm. In their system, flowers were classified into three categories of different structures: gamopetalous flowers, many-petaled flowers, and single-petaled flowers. For each structure, a classifier with specific feature set was constructed. Fuzzy c-means clustering algorithm was then used to determine the degree of membership of each image to each structure. The overall similarity is a linear combination of each individual similarity computed for each classifier with the weight being the degree of membership. The test database consists of 448 images from 112 species with 4 images per species. Experimental results have shown that the multiple-classifier approach outperforms any single-classifier approach. However, it is too rough a classification mechanism to classify flowers into three different categories according to the number of petals.

Note that the previous researchers extracted color and shape features from the whole image region or flower boundary, without specifically treating the color and shape characteristics of the pistil/stamen area. Thus, an interactive flower image recognition system, which extracts the color and shape features not only from the whole flower region but also from the pistil/stamen area, will be proposed to describe the characteristics of the flower images more precisely. First, a flower segmentation method is developed to segment the flower boundary with as fewer user interactions as possible. Further, a simple normalization procedure is employed to make the extracted features more robust to shape deformations, including the number of petals, the relative positions of petals, the poses of petals taken from different directions, flower sizes, etc. The rest of this paper is organized as follows. Section 2 describes the proposed flower image recognition system. Some experimental results will be given in Section 3. Conclusions will be given in Section 4.

2 The proposed flower image recognition system

The proposed flower image recognition system consists of three major phases: flower region segmentation, feature extraction, and recognition, as shown in Fig. 1. In the segmentation phase, the proposed system provides an interface allowing a user to draw a rectangular window which circumscribes the flower region. A segmentation algorithm similar to that proposed by Saitoh et al. [13] is then developed to extract the flower region within the rectangular window. In the feature extraction phase, the shape and color features of the whole flower region as well as the pistil/stamen area are extracted to measure the similarity between two flower images. In the recognition phase, the flower image in the database that is most similar to the input image will be found using the extracted features.

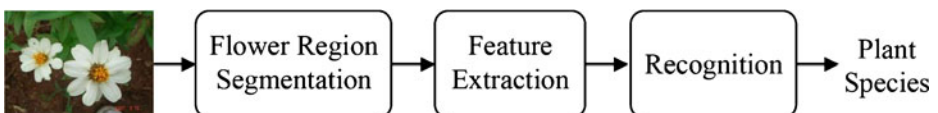


Fig. 1 Flow diagram of the proposed flower image recognition system

2.1 Flower region segmentation

In order to extract the flower boundary as correctly as possible, the proposed system provides a simple interactive interface which allows the user to select the interested flower for recognition. Figure 2 illustrates the steps of the interactive flower region segmentation phase. First, the user can draw a rectangular window which circumscribes the interested flower by using mouse click and drag operations. Let P_0 denote the center point of the rectangular window, P_1 , P_2 , P_3 , and P_4 denote the middle points on each of the four boundary lines of the rectangular window as shown in Fig. 3. For each scan line starting from any P_i ($i=1, 2, 3, 4$) to P_0 , the edge point locating at the flower boundary will be detected. These four edge points will then be regarded as the starting/ending points for boundary tracing. Since the proposed flower edge detection method used the “local cost” value associated with every pixel on each scan line, we will define the local cost first.

2.1.1 Definition of local cost

The local cost (LC) of a pixel on a scan line is defined as follows:

$$LC = 1 + MG - G, \quad (1)$$

where G denotes the gradient magnitude for the pixel and MG denotes the maximum gradient magnitude of all pixels in the image. According to the definition of local cost, we can see that a pixel with strong edge gradient will have a small local cost. In this paper, the Sobel operators (see Fig. 4) are employed to compute the horizontal and vertical gradient magnitudes of a pixel. Let $I_R(x, y)$, $I_G(x, y)$, and $I_B(x, y)$ denote respectively the R, G, and B color values of a pixel locating at (x, y) . For R color value, the corresponding gradient magnitude, notated $G_R(x, y)$, is defined as follows:

$$G_R(x, y) = \sqrt{G_{R,H}^2(x, y) + G_{R,V}^2(x, y)}, \quad (2)$$

where $G_{R,H}(x, y)$ and $G_{R,V}(x, y)$ denote respectively the horizontal and vertical gradients derived by convolving the horizontal and vertical Sobel operators with the 3×3 image blocks centered at (x, y) and can be described by the following equations:

$$\begin{aligned} G_{R,H}(x, y) = & [I_R(x + 1, y - 1) + 2I_R(x + 1, y) + I_R(x + 1, y + 1)] \\ & - [I_R(x - 1, y - 1) + 2I_R(x - 1, y) + I_R(x - 1, y + 1)] \end{aligned} \quad (3)$$

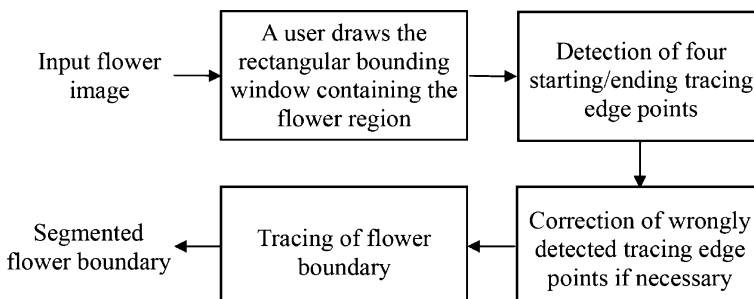
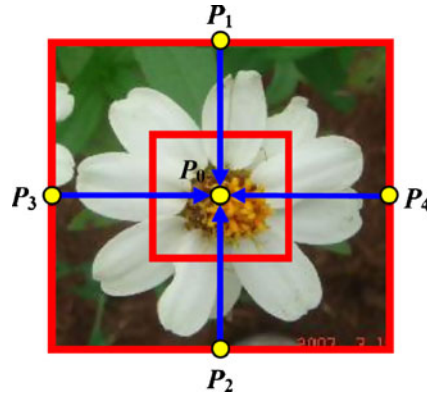


Fig. 2 Block diagram for flower region segmentation

Fig. 3 Four boundary lines on the flower bounding rectangular window



and

$$G_{R,V}(x,y) = [I_R(x - 1,y + 1) + 2I_R(x,y + 1) + I_R(x + 1,y + 1)] - [I_R(x - 1,y - 1) + 2I_R(x,y - 1) + I_R(x + 1,y - 1)] \quad (4)$$

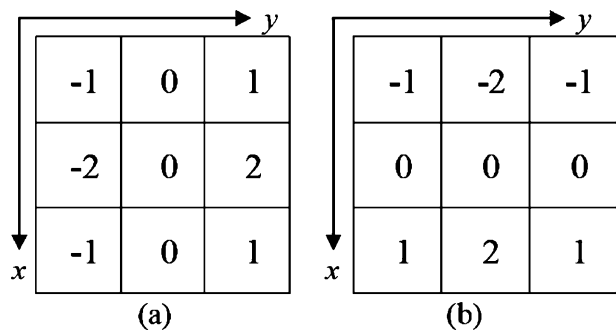
The gradient magnitudes for G and B color values, notated $G_G(x, y)$ and $G_B(x, y)$, can be computed in a similar manner. The overall gradient is then defined as the maximum value among $G_R(x, y)$, $G_G(x, y)$, and $G_B(x, y)$:

$$G(x,y) = \max \{G_R(x,y), G_G(x,y), G_B(x,y)\} \quad (5)$$

2.1.2 Detection of flower edge points

According to the computed local cost associated with each pixel, four profiles of local costs along the lines starting from every P_i ($i=1, 2, 3, 4$) to P_0 are generated. In this study, the estimated stamen region will be excluded on each profile (see $P_3 \rightarrow P_5$ in Fig. 5). The estimated stamen region is defined as the rectangular window with its center locating at P_0 and its area being 1/9 of the flower bounding window. For each profile, the 5th percentile, $P_{LC}(5)$, of local costs is evaluated. The threshold value, T_{LC} , used to find edge points on each profile is defined as the average of the local costs smaller than $P_{LC}(5)$. If the local cost value of a point is smaller than the threshold T_{LC} , it will be considered as a candidate edge point. The one closest to the border of the flower bounding window is regarded as the

Fig. 4 The Sobel operators **a** vertical Sobel operator **b** horizontal Sobel operator



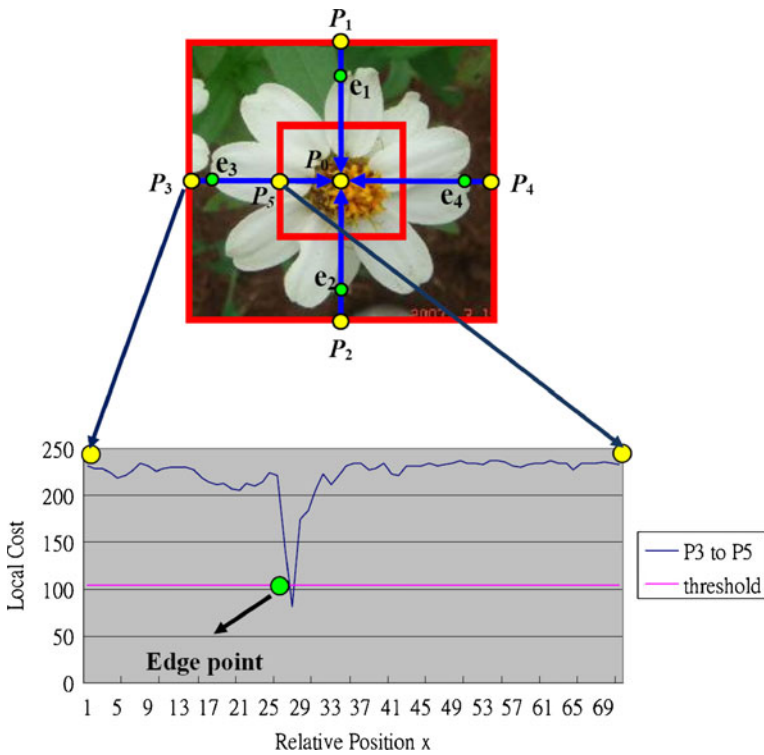


Fig. 5 An example for the detection of flower edge point on each profile

flower edge point (see e_1 , e_2 , e_3 and e_4 in Fig. 5). These four flower edge points will be taken as the starting/ending points of the flower boundary tracing algorithm. In our experiments, about 14.4% (50/348) among the 348 images in our database contain at least one wrongly detected edge points. Figure 6 gives some examples of wrongly detected edge points. The main reasons for producing these wrongly detected edge points are: 1) there exist strong edge points within the flower region (see Fig. 6(a)); 2) the contrast between the flower region and the background is not sharp enough (see Fig. 6(b)); 3) overlapping between neighboring flowers (see Fig. 6(c)); 4) no flower edge point survives in the profile when the stamen region is excluded (see Fig. 6(d)), etc. For these images, we provide an interactive interface which allows the user to use the mouse to select the correct edge point (see Fig. 7).

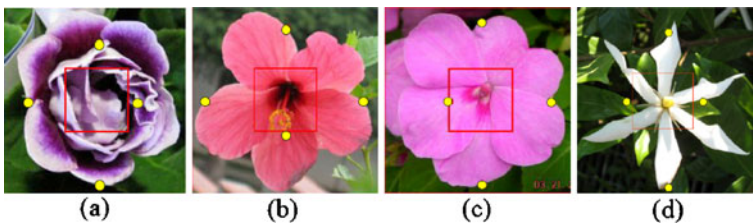


Fig. 6 Some example images showing that the edge points detected are incorrect

2.1.3 Flower boundary tracing

Let e_1 , e_2 , e_3 , and e_4 denote the detected flower edge points. The two lines which connect P_1 and P_2 as well as P_3 and P_4 will divide the flower bounding window into four sub-regions (see R_1 , R_2 , R_3 , and R_4 in Fig. 8). The flower boundary of each sub-region will be independently traced. Each pair of the four sets of edge points (e_1, e_4) , (e_4, e_2) , (e_2, e_3) and (e_3, e_1) will serve as the starting and ending tracing points of each sub-region, respectively. The proposed flower boundary tracing algorithm starts from the starting point and stops when the ending point is reached. These four partial flower boundaries will then be combined to form the whole flower boundary (see the yellow curve in Fig. 8).

The proposed flower boundary tracing algorithm modifies the 2-D dynamic programming graph search algorithm developed by Mortensen et al. [9]. It treats each pixel within the flower bounding window a vertex in a graph. An edge in the graph will connect a pixel to one of its 8-connected neighboring pixels. The cost associated with an edge is defined as the local cost evaluated on the neighboring pixel. The concept of average path cost, which is defined as the partial average cost computed from the previous pixel to the next pixel, is employed to decide which direction to move. The partial average cost is updated by adding the average of the previous pixel cost and the next pixel cost. The detailed algorithm of the modified flower boundary tracing algorithm is described as follows.

Algorithm: flower boundary tracing algorithm

Input:

s Starting pixel
 e Ending pixel
 $c(q, r)$ Cost function for the edge between pixels q and r

Output:

p Pointers from each pixel to its parent pixel with the minimum cost path

Data Structures:

L List of active pixels (i.e. pixels that are not determined with the minimum average cost yet and will be chosen as candidates to expand at next step) sorted by average cost (initially empty)
 $N(q)$ 8-neighboring pixels of q
 $B(q)$ Boolean function indicating if q has been expanded/processed
 $G(q)$ Cost function from the starting pixel s to q

Algorithm:

```

G(s) = 0; L = L + {s};           //Initialize the active list and the cost function
while (L ≠ ∅ and not B(e)) do begin //While there exists pixels to expand
  q ← min(L);                   //Find the pixel from L with minimum cost
  B(q) = TRUE;                  //Set q as expanded (i.e. processed)
  for each r ∈ N(q) with not B(r) do begin
    temp = (G(q)+c(q, r))/2;     //Compute average cost to r
    if r ∈ L and temp < G(r) then
      L = L - {r};              //Remove higher cost pixel r from L
    if r ∉ L then begin          //If r is not in L
      G(r) = temp;              //update average cost for r
      p(r) = q;                 //set the parent pixel of r
      L = L + {r};              //and place r in L
    end
  end
end
end

```

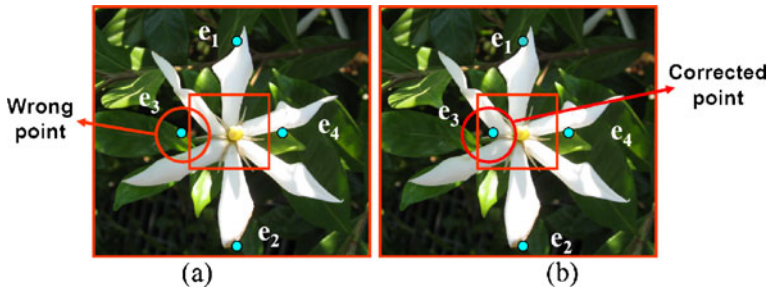



Fig. 7 Correction of a wrongly detected edge point **a** a wrong edge point, e_3 **b** The corrected edge point through user interaction

2.2 Feature extraction

The most widely used features for describing flowers are color and shape descriptors. In this paper, the color and shape features of the whole flower region and the pistil/stamen area will be extracted in an attempt to describe the characteristics of the flower images more precisely.

2.2.1 Features of the whole flower region

First, we define the flower region as the internal region within the segmented flower boundary. In this paper, nine color features in which the first six color features were proposed by Saitoh et al. [13] and three shape features are extracted from the whole flower region for recognition purpose.

Color features of flower region Since the flower images were taken in different environmental conditions, the variation in illumination will greatly affect the recognition result. To deal with such a problem, we convert each pixel from the *RGB* color space to *HSV* (hue, saturation, and value) space [6] and discard the illumination (*V*) component. The color features are derived from the primary, secondary, and thirdly flower colors appearing in the whole flower region. First, the *HS* space is divided into 12×6 color cells represented by C_i , $1 \leq i \leq 72$ (please see Fig. 9). The color coordinate of each cell, which is defined as the coordinate of the center point of each cell, can be represented by a pair of *H* and *S* values, (H_i, S_i) , $1 \leq i \leq 72$. For each flower region, a color histogram (notated $CH(i)$, $1 \leq i \leq 72$), which describes the probability associated with each color cell C_i , will be generated. Let $DC(1)$, $DC(2)$, and $DC(3)$ denote respectively the first three dominant color cells

Fig. 8 The flower bounding window is divided into four sub-regions

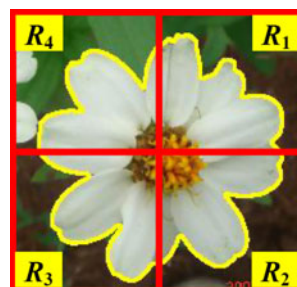
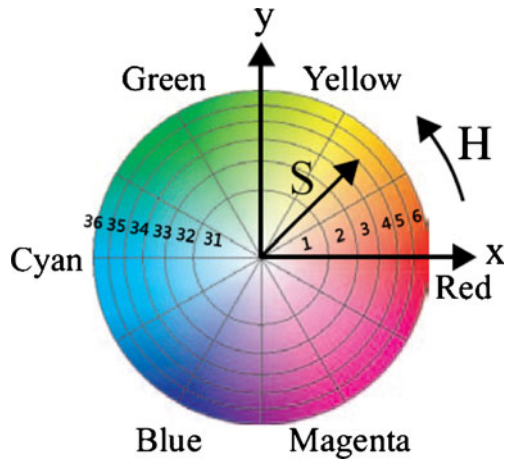


Fig. 9 The *HS* color space is divided into 12×6 color cells



appearing in the flower region. The color coordinates of these three dominant color cells and their corresponding probabilities are taken as the color features of the flower region. Let (dx_i, dy_i) and p_i denote the coordinate vector and the corresponding probability of $DC(i)$, $1 \leq i \leq 3$, where $dx_i = S_{DC(i)} \cos(H_{DC(i)})$ and $dy_i = S_{DC(i)} \sin(H_{DC(i)})$. These color features can be summarized as follows.

- CF_1 : x -coordinate value of DC_1 , dx_1
- CF_2 : y -coordinate value of DC_1 , dy_1
- CF_3 : the probability of DC_1 , p_1
- CF_4 : x -coordinate value of DC_2 , dx_2
- CF_5 : y -coordinate value of DC_2 , dy_2
- CF_6 : the probability of DC_2 , p_2
- CF_7 : x -coordinate value of DC_3 , dx_3
- CF_8 : y -coordinate value of DC_3 , dy_3
- CF_9 : the probability of DC_3 , p_3

Shape features of flower region To get the shape features, we first define the centroid (g_x, g_y) of the flower region as the flower center, which is computed as follows:

$$g_x = \frac{1}{N} \sum_{i=1}^N x_i, \tag{6}$$

$$g_y = \frac{1}{N} \sum_{i=1}^N y_i, \tag{7}$$

where N is the number of pixels on the flower boundary, x_i and y_i are respectively the x and y coordinates of the i -th boundary pixel. The distance between the flower center and each boundary pixel is then computed as follows:

$$d_i = \sqrt{(x_i - g_x)^2 + (y_i - g_y)^2}, \quad 1 \leq i \leq N. \tag{8}$$

Without loss of generality, let the distances be sorted in an increasing order. That is, $d_i \leq d_{i+1}$, for $1 \leq i \leq N-1$. The three shape features (notated SF_1 , SF_2 , and SF_3) used to represent the shape characteristics of the flower region will be described as follows.

- 1) SF_1 : A ratio which indicates the relevant sharpness of the petals and is computed from the distances between the flower boundary points to the flower center defined as follows:

$$SF_1 = \frac{R_{10}}{R_{90}}, \tag{9}$$

where R_{10} and R_{90} are respectively the average distances among all d_i 's which are smaller than 10th percentile and larger than 90th percentile of all d_i 's:

$$R_{10} = \frac{1}{0.1 \times N} \sum_{i=1}^{0.1 \times N} d_i, \tag{10}$$

$$R_{90} = \frac{1}{0.1 \times N} \sum_{i=1}^{0.1 \times N} d_{N-i}. \tag{11}$$

Note that the computed feature value SF_1 defined as a ratio between R_{10} and R_{90} will not change greatly when the flower region is broken or captured from different directions.

- 2) SF_2 : The average of normalized distances computed from every flower boundary point to the flower center defined as follows:

$$SF_2 = \frac{1}{N} \sum_{i=1}^N D_i, \tag{12}$$

where D_i is the normalized distance defined as follows:

$$D_i = \begin{cases} 1, & d_i \geq R_{90} \\ \frac{d_i - R_{10}}{R_{90} - R_{10}}, & R_{10} < d_i < R_{90} \\ 0, & d_i \leq R_{10} \end{cases}. \tag{13}$$

Note that the definition of the feature value SF_2 using the averaged normalized values D_i 's will make it invariant to the size of the flower region.

- 3) SF_3 : Roundness measure which indicates how much the shape of the flower petal is closer to a circle and is defined as follows:

$$SF_3 = \frac{4\pi S}{L^2}, \tag{14}$$

where L is the length of the flower boundary and S is the area of the flower region defined as the total number of pixels in the flower region. When the flower shape is close to a circle, SF_3 will be close to 1. Note that this feature value is robust to rotation, translation, and scaling of flower objects.

In summary, the 12-dimensional feature vector used to represent the flower region can be described as follows:

$$\mathbf{f}_F = [CF_1, CF_2, \dots, CF_9, SF_1, SF_2, SF_3]^T. \tag{15}$$

2.2.2 Features of the pistil/stamen area

First, we define an initial estimate of the pistil/stamen area as the square area with its center locating at the flower center and its width $2/3$ of the petal length, where the petal length is defined as R_{90} in (11). Let $PDC(1)$ denote the dominant color cell in this estimated pistil/stamen area. Note that $PDC(1)$ is found by excluding the primary color appearing in the flower region, $DC(1)$. Then, all image pixels within the square area of width $4/3$ of the petal length having color values identical to that of $PDC(1)$ will constitute the pistil/stamen area.

Since the color and shape of the pistil/stamen area also exhibit some discriminating information for flower image recognition, the dominant color and its corresponding probability will be taken as the color features of the pistil/stamen area. In addition, the mean, standard deviation, and the third central moment of the normalized distance from each pixel in the pistil/stamen area to the center of the pistil/stamen area will be computed as the shape features of the pistil/stamen area.

Color features of pistil/stamen area For most flowers, the dominant color of the pistil/stamen area is often different from that of the flower region. Thus, the color characteristic of the pistil/stamen area provides some discriminating information. In this study, the coordinate vector (pdx_1, pdy_1) and the corresponding probability pp_1 of $PDC(1)$ will be taken as the color features of the pistil/stamen area, where $pdx_1 = S_{PDC(1)} \cos(H_{PDC(1)})$ and $pdy_1 = S_{PDC(1)} \sin(H_{PDC(1)})$. These color features can be summarized as follows.

- PCF_1 : x -coordinate value of PDC_1 , pdx_1
- PCF_2 : y -coordinate value of PDC_1 , pdy_1
- PCF_3 : the probability of PDC_1 , pp_1

Shape features of pistil/stamen area Let the pistil/stamen area consist of M pixels and the coordinates of these M pixels be notated as (px_i, py_i) , $1 \leq i \leq M$. Next, the centroid (g_{px}, g_{py}) of the pistil/stamen area is computed as follows:

$$g_{px} = \frac{1}{M} \sum_{i=1}^M px_i, \tag{16}$$

$$g_{py} = \frac{1}{M} \sum_{i=1}^M py_i. \tag{17}$$

The distance between the centroid and every pixel of the pistil/stamen area is then computed as follows:

$$pd_i = \sqrt{(px_i - g_{px})^2 + (py_i - g_{py})^2}, \quad 1 \leq i \leq M. \tag{18}$$

Without loss of generality, let the distances be sorted in an increasing order. That is, $pd_i \leq pd_{i+1}$, $1 \leq i \leq M-1$. These distances are then normalized using the following equation:

$$PD_i = \begin{cases} 1, & pd_i \geq PR_{90} \\ \frac{pd_i - PR_{10}}{PR_{90} - PR_{10}}, & PR_{10} < pd_i < PR_{90} \\ 0, & pd_i \leq PR_{10} \end{cases}, \tag{19}$$

where PR_{10} and PR_{90} are respectively the average distances among all pd_i 's which are smaller than 10th percentile and larger than 90th percentile of all pd_i 's:

$$PR_{10} = \frac{1}{0.1 \times M} \sum_{i=1}^{0.1 \times M} pd_i, \quad (20)$$

$$PR_{90} = \frac{1}{0.1 \times M} \sum_{i=1}^{0.1 \times M} pd_{M-i}. \quad (21)$$

Note that the normalized distance values PD_i 's will make the extracted shape features of the pistil/stamen area invariant to the size of the flower image. The three shape features (notated PSF_1 , PSF_2 , and PSF_3) used to represent the shape characteristics of the pistil/stamen area are defined as follows.

1) PSF_1 : The mean of the normalized distance values PD_i 's, defined as follows:

$$PSF_1 = \mu_{PD} = \frac{1}{M} \sum_{i=1}^M PD_i. \quad (22)$$

2) PSF_2 : The standard deviation of the normalized distance values PD_i 's, defined as follows:

$$PSF_2 = \sigma_{PD} = \left(\frac{1}{M} \sum_{i=1}^M (PD_i - \sigma_{PD})^2 \right)^{1/2}. \quad (23)$$

3) PSF_3 : The third central moment of the normalized distance values PD_i 's, defined as follows:

$$PSF_3 = m_3 = \left(\frac{1}{M} \sum_{i=1}^M (PD_i - \sigma_{PD})^3 \right)^{1/3}. \quad (24)$$

Note that these three shape features are typically robust to rotation, translation, and scaling of flower images. In summary, the 6-dimensional feature vector used to represent the pistil/stamen area can be described as follows:

$$\mathbf{f}_P = [PCF_1, PCF_2, PCF_3, PSF_1, PSF_2, PSF_3]^T. \quad (25)$$

The feature descriptor used to represent a flower image consists of all 18 features extracted from the flower region as well as the pistil/stamen area:

$$\begin{aligned} \mathbf{f} &= [f(1), f(2), \dots, f(18)]^T = \left[(\mathbf{f}_F)^T (\mathbf{f}_P)^T \right]^T \\ &= [CF_1, \dots, CF_9, SF_1, \dots, SF_3, PCF_1, \dots, PCF_3, PSF_1, \dots, PSF_3]^T. \end{aligned} \quad (26)$$

2.3 Flower image recognition

In the recognition phase, the distances between the input image and all flower images in the database are calculated. The distance between the input image and the i -th image

in the database, notated $dist_i$, is measured by the weighted Euclidean distance defined as follows:

$$dist_i = \frac{\sum_{k=1}^{N_f} w(k) |f_i(k) - f(k)|}{\sum_{k=1}^{N_f} w(k)}, \tag{27}$$

where $f_i(k)$ is the k -th feature value of the i -th image, $f(k)$ is the k -th feature value of the input image, and $w(k)$ is the weight associated with the k -th feature value. The variable N_f determines which type of features is used for flower image recognition. If only the features extracted from the flower region are used for recognition purpose, set $N_f=12$; if features extracted from both the flower region and the pistil/stamen area are used for recognition purpose, set $N_f=18$. In this paper, the normalized top five recognition rate, $NAR5(k)$, is used as the weight associated with the k -th feature value ($1 \leq k \leq 18$). That is, $w(k) = NAR5(k)$. $NAR5(k)$ is derived by normalizing the top five recognition rate associated with the k -th feature value, $AR5(k)$:

$$w(k) = NAR5(k) = \frac{AR5(k)}{\sum_{j=1}^{N_f} AR5(j)} \times 100, \tag{28}$$

where $AR5(k)$ denotes the recognition rate when the k -th feature value is individually used for flower recognition and the recognition result will be regarded as accurate if at least one of the species in the top five candidates is identical to the input one. Table 1 shows the top five recognition rate ($AR5(k)$) and the weight ($w(k)$) associated with each feature value for the first flower image database. Note that these initial weights are distinct for different databases. Based

Table 1 Top five recognition rate $AR5(k)$ and weight $w(k)$ associated with each feature value for the first flower image database

Feature	$AR5(k)$ (%)	$w(k)$
$f(1) : CF_1$	66.67	6.91
$f(2) : CF_2$	68.39	7.08
$f(3) : CF_3$	58.62	6.07
$f(4) : CF_4$	63.51	6.58
$f(5) : CF_5$	66.95	6.93
$f(6) : CF_6$	43.97	4.55
$f(7) : CF_7$	61.49	6.37
$f(8) : CF_8$	57.47	5.95
$f(9) : CF_9$	45.11	4.67
$f(10) : SF_1$	58.05	6.01
$f(11) : SF_2$	39.94	4.14
$f(12) : SF_3$	60.63	6.28
$f(13) : PCF_1$	59.77	6.19
$f(14) : PCF_2$	59.77	6.19
$f(15) : PCF_3$	54.31	5.63
$f(16) : PSF_1$	34.48	3.57
$f(17) : PSF_2$	29.60	3.07
$f(18) : PSF_3$	36.78	3.81

on the computation of all weighted distances, the top K candidate images with minimum distances to the input image will be returned to the user. In this study, the top K recognition rate will be employed to evaluate the recognition performance.

3 Experimental results

3.1 Flower image databases

In this paper, two flower image databases are used to evaluate the performance of the proposed method. The first flower image database constructed by us consists of 348 flower images from 24 species. A summary of these 24 plants is shown in Table 2 and some example images are shown in Fig. 10. The second database constructed by Zou and Nagy [16] consists of 612 flower images from 102 species.

The flower images in the first database were taken in the fields by using digital cameras equipped with macro lens. The range of the lens aperture specified as an F-number was set between F2.8 and F5.6. To ensure the robustness of the proposed system, a number of different cameras, including SONY T9, Canon IS 860, and NIKON S1, were used to take

Table 2 Common names and Latin names of the plant species and their corresponding image numbers in the first database

SC	Common names	Latin names	Image numbers
1	Yellow dots	Wedelia trilobata (L.) Hitchc.	14
2	Globe amaranth	Gomphrena globosa L.	8
3	African Marigold	Tagetes erecta Linn.	6
4	Zinnia	Zinnia elegans Jacq.	16
5	French Marigold	Tagetes patula L.	4
6	Sunflower	Helianthus annuus Linn.	6
7	Golden trumpet	Allamanda cathartica L.	12
8	Common Cosmos	Cosmos bipinnatus Cav.	9
9	Chinese hibiscus	Hibiscus rosa-sinensis L.	17
10	Purple Glory Tree	Tibouchina semidecandra	19
11	Bojers Spurge	Euphorbia milii Ch. des Moulins	9
12	Pilose Beggarticks	<i>Bidens pilosa</i> L. var. <i>radiata</i> Sch.	27
13	Pericuinkle1	Catharathus roseus (L.) Don	10
14	Phalaenopsis	Phalaenopsis amabilis	8
15	Marguerite	Chrysanthemum frutescens	17
16	Pericuinkle2	Catharathus roseus (L.) Don	6
17	Gloxinia	Gloxinia × hybrida Hort.	11
18	Anthurium	Anthurium scherzerianum Schott	13
19	Gardenia	Gardenia jasminoides Ellis	10
20	Primrose Jasmine	Jasminum mesnyi Hance	33
21	Lavender Sorrel	Oxalis martiana Zucc.	28
22	Putchock Fringed iris Crested iris	Iris japonica	18
23	Azalea	Rhododendron spp.	15
24	African Touch-me-not	Impatiens walleriana	32



Fig. 10 Some example images in the first database

these flower images. In addition, the images of the same species were taken from different flowers with different poses. The number of images in each species is different and ranges from 4 to 33. Several images contain multiple, tiny, overlapping flowers as shown in Fig. 10. Before recognition, all flower images were re-scaled to be the same size of 400×300 pixels.

The second database, which was constructed by Zou and Nagy, consists of 612 flower images from 102 species collected from (<http://www.ecse.rpi.edu/doclab/flowers>). Each species consists of six images. The size of these images are identical, 300×240 pixels. Some pictures are quite out of focus, and several pictures contain multiple, tiny, overlapping flowers.

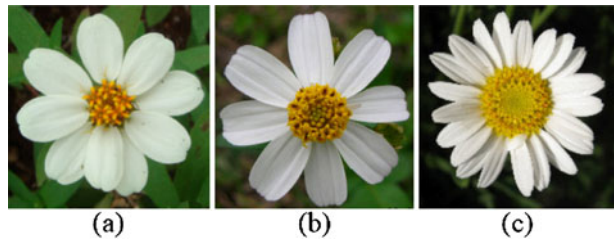
3.2 Recognition results

To show the effectiveness of the proposed approach, we will compare the recognition results of the proposed approach with the feature sets proposed by Hong et al. [7], Zou and

Table 3 Comparison of recognition rate on the first flower image database

Method	Recognition rate (%)				
	Top 1	Top 2	Top 3	Top 4	Top 5
Hong et al. [7]	51.4	64.4	76.2	81.3	85.6
Saitoh et al. [13]	83.6	92.0	95.1	97.7	99.1
Our proposed approach 1 (using only shape features of flower region)	47.1	64.9	78.2	82.5	87.4
Our proposed approach 2 (using only color features of flower region)	77.9	88.8	92.5	94.8	96.3
Our proposed approach 3 (using color and shape features of flower region)	86.2	93.4	96.3	97.4	98.3
Our proposed approach 4 (using color and shape features of flower region and shape features of pistil/stamen area)	87.1	94.0	96.6	97.4	98.3
Our proposed approach 5 (using color and shape features of flower region and color features of pistil/stamen area)	88.5	95.1	98.6	99.7	99.7
Our proposed approach 6 (using color and shape features of flower region and pistil/stamen area)	89.1	95.1	98.6	99.7	99.7

Fig. 11 The 3 different species with similar appearance **a** Zinnia **b** Pilose Beggarticks **c** Marguerite



Nagy [16], and Saitoh et al. [13]. The overall recognition accuracy is evaluated by taking each flower image in the database as an input image and the remaining flower images are considered as the training set.

For the first flower image database, the comparison of the recognition rate of the proposed approach with those proposed by Hong et al. [7] and Saitoh et al. [13] is shown in Table 3. We can see that the color features give more discriminating ability than the shape features. The proposed approach using both shape and color features of the flower region can achieve a very high recognition rate. Furthermore, the best performance is obtained by combining the shape and color features of both the flower region as well as the pistil/stamen area, which outperforms the approaches proposed by Hong et al. [7] and Saitoh et al. [13] in terms of the recognition rate. Figure 11 shows in the first flower image database the three different species (Zinnia, Pilose Beggarticks, and Marguerite) with similar appearance. The comparison of the recognition rate on these species is shown in Table 4. From this table, we can see that our proposed approach also outperforms the other two approaches. Specifically, Fig. 12 shows the 27 images of Pilose Beggarticks. We can see that these images exhibit different characteristics, including the number of petals, the relative positions of petals, the shape of petals, etc. Furthermore, for some images there are some bugs appearing in the different part of the flower region. From Table 4, we can see the comparison of recognition rates on this species. It is clear that the proposed approach yields the best recognition rate. This result illustrates that the proposed approach is more robust to noises and shape variations.

For the second database, we have also compared the recognition rate of the proposed approach with those proposed by Hong et al. [7], Zou and Nagy [16], and Saitoh et al. [13].

Table 4 Comparison of recognition rate on the three species (Zinnia, Pilose Beggarticks, and Marguerite)

Plant species	Method	Recognition rate (%)				
		Top 1	Top 2	Top 3	Top 4	Top 5
Zinnia	Hong et al. [7]	50	50	50	50	50
	Saitoh et al. [13]	50	50	50	75	100
	Our proposed approach 6	100	100	100	100	100
Pilose Beggarticks	Hong et al. [7]	18.5	44.4	85.2	85.2	92.6
	Saitoh et al. [13]	77.8	92.6	96.3	100	100
	Our proposed approach 6	88.9	100	100	100	100
Marguerite	Hong et al. [7]	35.3	52.9	70.6	82.4	88.2
	Saitoh et al. [13]	70.6	70.6	88.2	94.1	100
	Our proposed approach 6	82.4	94.1	100	100	100



Fig. 12 The 27 images of Pilose Beggarticks in the first image database

Table 5 compares the recognition rates of these approaches. From the table, we can see that our proposed approach using shape and color features of the flower region and the pistil/stamen area always yields a higher recognition rate than the other approaches. Note that the method proposed by Zou and Nagy achieved a recognition rate of 93% when a number of user interactions were repeatedly conducted until the user accepted the recognition result, which takes a longer computation time (10.7 s) than our proposed approach (4.2 s). Figure 13 shows the images of two flower species (FlowerBridge1 and RosaArkansana) in the second database. We can see that the images of the same species were taken from different directions and distances and thus these images reveal different shapes and sizes. The comparison of the recognition rate on these two species is shown in Table 6. From this table, we can see that our proposed approach also outperforms the other methods.

Table 5 Comparison of recognition rate on the second flower image database

Method	Recognition rate (%)					Average recognition time (s)
	Top 1	Top 2	Top 3	Top 4	Top 5	
Hong et al. [7]	39.5	50.8	57.0	61.1	64.4	4.2
Zou-Nagy et al. [16] (without user interaction)	52	–	79	–	–	8.5
Zou-Nagy et al. [16] (with user interaction)	93					10.7
Saitoh et al. [13]	65.5	81.9	87.9	91.5	93.8	4.2
Our proposed approach 1	12.8	22.6	28.4	36.1	43.5	4.2
Our proposed approach 2	56.7	72.9	80.2	85.3	88.1	4.2
Our proposed approach 3	69.3	84.0	89.1	92.7	94.8	4.2
Our proposed approach 4	71.3	84.6	89.9	92.8	95.1	4.2
Our proposed approach 5	76.1	88.7	95.9	97.2	98.2	4.2
Our proposed approach 6	77.8	89.4	94.3	97.7	98.5	4.2

Fig. 13 The images of the two species in the second database **a** FlowerBridge1 **b** RosaArkansana



4 Conclusions

In this paper, we have presented an interactive flower recognition system. First, the system provides a simple user interface which allows each user to draw a rectangular bounding window containing the interested flower region. Then, a boundary tracing algorithm is developed to find the flower boundary as accurately as possible. Besides the color and shape features of the whole flower region, the color and shape features of the pistil/stamen area are also extracted to represent the flower characteristics in a more precise fashion. Experiments conducted on two different flower image databases consisting of 24 species and 102 species have shown that our proposed approach achieves a higher recognition rate than the methods proposed by Hong et al. [7], Zou and Nagy [16], and Saitoh et al. [13].

Table 6 Comparison of recognition rate on FlowerBridge1 and RosaArkansana in the second database

Plant species	Method	Recognition rate (%)				
		Top 1	Top 2	Top 3	Top 4	Top 5
FlowerBridge1	Hong et al. [7]	33.3	33.3	33.3	33.3	33.3
	Saitoh et al. [13]	33.3	50.0	50.0	83.3	100
	Our proposed approach 6	66.7	83.3	100	100	100
RosaArkansana	Hong et al. [7]	0.0	33.3	50.0	50.0	50.0
	Saitoh et al. [13]	33.3	66.7	66.7	66.7	66.7
	Our proposed approach 6	100	100	100	100	100

References

1. Boykov YY, Jolly MP (2001) Interactive graph cuts for optimal boundary and region segmentation of objects in N-D images. *Proc Int Conf Comput Vis* 2:105–112
2. Cho SY, Chi Z (2005) Generic evolution processing of data structures for image classification. *IEEE Trans Knowl Data Eng* 17(2):216–231
3. Das M, Manmatha R, Riseman EM (1999) Indexing flower patent images using domain knowledge. *IEEE Intell Syst* 14:24–33
4. Fukuda K, Takiguchi T, Arika Y (2008) Multiple classifier based on fuzzy c-means for a flower image retrieval. In: *Proc. Int. Workshop on Nonlinear Circuits and Signal Processing (NCSP'2008)*: 76–79
5. Goller C, Kuchler A (1996) Learning task-dependent distributed representations by back-propagation through structure. *Proc IEEE Int Conf Neural Networks* 1:347–352
6. Gonzalez RC, Woods RE (2002) *Digital image processing*, 2nd edn. Prentice-Hall, Upper Saddle River
7. Hong A, Chen G, Li J, Chi Z, Zhang D (2004) A flower image retrieval method based on ROI feature. *Journal of Zhejiang University (Science)* 5(7):764–772
8. Lowe D (2004) Distinctive image features from scale-invariant keypoints. *Int J Comput Vis* 60(2):91–110
9. Mortensen EN, Barrett WA (1995) Intelligent scissors for image composition. In: *Proc. Computer Graphics (SIGGRAPH'95)*: 191–198
10. Nilsback ME, Zisserman A (2006) A visual vocabulary for flower classification. *Proc Comput Vis Pattern Recogn* 2:1447–1454
11. Saitoh T, Kaneko T (2000) Automatic recognition of wild flowers. *Proc Int Conf Pattern Recogn* 2:507–510
12. Saitoh T, Aoki K, Kaneko T (2003) Automatic extraction of object region from photographs. In: *Proc. Scandinavian Conf. on Image Analysis*: 1130–1137
13. Saitoh T, Aoki K, Kaneko T (2004) Automatic recognition of blooming flowers. *Proc Int Conf Pattern Recogn* 1:27–30
14. Swain MJ, Ballard DH (1991) Color indexing. *Int J Comput Vis* 7(1):11–32
15. Varma M, Zisserman A (2002) Classifying images of materials: achieving viewpoint and illumination independence. *Proc Europ Conf Comput Vis* 3:255–271
16. Zou J, Nagy G (2004) Evaluation of model-based interactive flower recognition. *Proc Int Conf Pattern Recogn* 2:311–314



Tzu-Hsiang Hsu was born on Feb 7, 1984 in Taoyuan, Taiwan. She received the M.S. degree in Multimedia Engineering from National Chiao Tung University, Hsinchu, Taiwan in 2008. She is currently an engineer of the software development department of ZyXEL, a network terminal product company. Her research interests include image processing, pattern recognition and segmentation.



Chang-Hsing Lee was born on July 24, 1968 in Tainan, Taiwan. He received the B.S. and Ph.D. degrees in Computer and Information Science from National Chiao Tung University, Hsinchu, Taiwan in 1991 and 1995, respectively. He is currently an Associate Professor in the Department of Computer Science and Information Engineering, Chung Hua University, Hsinchu, Taiwan. His main research interests include audio/sound classification, multimedia information retrieval, and image processing.



Ling-Hwei Chen received the B.S. degree in mathematics and the M.S. degree in applied mathematics from National Tsing Hua University, Hsinchu, Taiwan in 1975 and 1977, respectively, and the Ph.D. in computer engineering from National Chiao Tung University, Hsinchu, Taiwan in 1987